LA-UR-20-26488
Approved for public release; distribution is unlimited.

Title: How data science methods can improve the quality and efficiency of ICF and HEDP research

Author(s): Knapp, Patrick F.
Glinsky, Michael E.
Tobias, Benjamin John
Kline, John L.

Intended for: Report

Issued: 2020-08-24

How data science methods can improve the quality and efficiency of ICF and HEDP research

Patrick F. Knapp[1], Michael E. Glinsky[1], Ben Tobias[2], John Kline[2]
[1] *Sandia National Laboratories, Albuquerque NM, 87185*
[2] *Los Alamos National Laboratory, Los Alamos NM, 87545*

June 2020

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

CONTENTS

LIST OF FIGURES

EXECUTIVE SUMMARY

Data Science methods (many that are Bayesian based) are widely used in the physical sciences to estimate model parameters from experimental data, synthesize heterogeneous data, calibrate models, design experiments, and determine statistical significance of data. These methods provide a wealth of advantages over traditional analysis techniques because: 1) uncertainties are rigorously defined and propagated naturally through complex systems including covariance, 2) prior information is captured within the analysis framework (including rad-MHD and rad-hydro simulations), 3) competing models can be selected and/or ruled out using quantitative criteria, and 4) complex, heterogeneous data can be incorporated simultaneously. While these methods have been widely adopted as the gold standard in fields such as particle physics, astronomy, and biology, they have been slow to catch on in Inertial Confinement Fusion (ICF) and High Energy Density Physics (HEDP) research. Recently, several teams at LLNL, SNL, LANL, and the LLE have been exploring the use of these tools in their research and have found success. Here we propose that a concerted effort to consolidate these independent research efforts by developing and deploying common tools for use across the complex can revolutionize the way we approach data analysis, assimilation of theory and experiment, and decision making. The Bayesian formalism provides a means to accomplish this, but we are lacking certain infrastructure to make it happen on a large scale. Furthermore, once adopted, these techniques can be used to develop standards by which discoveries can be judged, similar to the so-called 5σ rule in high energy particle physics. Such standards may be used in the future to address the issue of unknown reproducibility in ICF and HED experiments caused by low shot rate and high cost per experiment. Our goals as a group are to advance the state of the art in HED measurement science by enabling: 1) better inferences from data with well-defined uncertainties, 2) better use of the data we have and continue to collect, 3) intelligent synthesis of data, 4) evaluation of the statistical significance of our data, and 5) informed decision making regarding the design of new experiments and instruments.

ACRONYMS AND DEFINITIONS

Abbreviation	Definition
SNL	Sandia National Laboratories
LANL	Los Alamos National Laboratory
LLNL	Lawrence Livermore National Laboratory
LLE	Laboratory for Laser Energetics (at the University of Rochester)
MLDL	Machine Learning/Deep Learning
DNN	Deep Neural Network
ICF	Inertial Confinement Fusion
HEDP	High Energy Density Physics
MHD	Magneto-Hydrodynamics
MCMC	Markov Chain Monte Carlo
NNSA	National Nuclear Security Administration
HPC	High Performance Computing
BIE	Bayesian Inference Engine, a tool developed at LANL
NIF	National Ignition Facility
OMEGA	Laser Facility located at the LLE

1. BACKGROUND AND INTRODUCTION

The application of Bayes' Theorem to the analysis of data in the physical sciences allows for the rigorous estimation of parameters with uncertainties in high-dimensional spaces (1) (2). Traditional methods (e.g. least squares fitting) do not generalize well to high dimensions or to the incorporation of large, heterogeneous data sets. This is exactly the space we find ourselves in HEDP, where we have a small number of experiments that employ a wide variety of diagnostics measuring different aspects of the complicated integrated physics of the experiment. Bayesian parameter estimation provides us with a cohesive framework with which to integrate these data using a physical model of the experiment. However, this process provides much more than just estimates of the best fit parameters. At their heart, Bayesian methods determine the *posterior probability distribution* of the model parameters. This quantity allows us to measure how likely any given combination of parameters is, providing rigorously defined confidence intervals with covariances between all model parameters as well as the measurements. Additionally, our prior knowledge of the physics of the problem and parameter space can be incorporated intuitively providing a regularization that enables solution of even ill posed problems.

With Bayesian inference, the basic problem we wish to solve is this: find the set of model parameters, $\overline{\boldsymbol{m}}$, that best matches a set of observables, $\overline{\boldsymbol{x}}$, with associated uncertainties, $\overline{\boldsymbol{\sigma}}$. Bayesian statistics provides us with a formalism that allows us to accomplish this task in a rigorous and quantitative way. In the Bayesian worldview the model can be viewed as a hypothesis about the physics describing the system. We have some degree of certainty that this hypothesis accurately describes our system given some prior knowledge about the parameters. This can be written mathematically as $\mathcal{P}(\overline{\boldsymbol{m}}\,|A)$, where A encapsulates our background assumptions about the system. Additionally, we view the set of model parameters, $\overline{\boldsymbol{m}}$, as having a continuum of values that describe the system with some *probability*, not as a set of correct values that we can only know with some finite certainty. In this interpretation, we wish to know the most probable set of $\overline{\boldsymbol{m}}$ by evaluating the probability distribution of agreement with the observations over the entire space of $\overline{\boldsymbol{m}}$. This function is written mathematically as $\mathcal{P}(\overline{\boldsymbol{x}}\mid\overline{\boldsymbol{m}},A)$, and is known as the *likelihood*. Formally, it is the probability of observing the data, $\overline{\boldsymbol{x}}$, given a set of model parameters, $\overline{\boldsymbol{m}}$, and our background assumptions A.

We have an expression for the probability of observing $\overline{\boldsymbol{x}}$ given a particular $\overline{\boldsymbol{m}}$ (the *likelihood*), and an expression encapsulating our prior assumptions about and knowledge of the system. What we want is an expression giving us the probability of a particular $\overline{\boldsymbol{m}}$, given our observations and assumptions. Bayes' Theorem gives us exactly that.

$$\mathcal{P}(\overline{\boldsymbol{m}}\mid\overline{\boldsymbol{x}},A)=\frac{\mathcal{P}(\overline{\boldsymbol{x}}\mid\overline{\boldsymbol{m}},A)\mathcal{P}(\overline{\boldsymbol{m}}\,|A)}{\mathcal{P}(\overline{\boldsymbol{x}}\,|A)}$$

where the term on the left hand side (known as the *posterior* probability) is precisely what we desire. The term in the denominator is known as the *evidence* and is generally simply a normalization constant (this is true for parameter estimation problems, but when performing model selection, this term is essential). A fundamental tenet of Bayesian inference is that it is impossible to analyze data without making assumptions. While this is always true, it is often not acknowledged or stated explicitly in practice. The Bayesian formalism requires that one write down these assumptions and explicitly define their relationship with the data, embodied in the term A above and the model being used.

The likelihood is determined by the problem at hand. For example, if the problem is to estimate the number of counts measured on a radiation detector, then one would use the Poisson distribution as their likelihood function. If the problem is to estimate the probability of flipping N heads over the course of M flips, then one would use the binomial distribution. In our case, and indeed in most data analysis applications, we use the normal distribution as our likelihood. It can be expressed as

$$\mathcal{P}(\overline{\boldsymbol{x}} \mid \overline{\boldsymbol{m}}, A) \propto \prod_{i=1}^{N} exp\left(-\frac{\mathcal{F}_i(\overline{\boldsymbol{m}}) - x_i}{2\sigma_i^2}\right)$$

This expression states that the likelihood is maximized where the function $\mathcal{F}_i(\overline{\boldsymbol{m}})$ equals x_i, and drops off with a standard deviation equal to the uncertainty in the data, σ_i. The functions $\mathcal{F}_i(\overline{\boldsymbol{m}})$ are the forward models that describe the mapping from model parameters to observations. In our case the x_i are the diagnostic measurements and the $\mathcal{F}_i(\overline{\boldsymbol{m}})$ are the synthetic diagnostic models used to calculate the measurements from a given configuration. The exponent in equation. Y is easily recognized as the familiar χ^2 for comparing a measurement to a model prediction.

Performing parameter estimation amounts to evaluating this equation by successive evaluations of the likelihood and prior probability (ignoring the evidence term for the moment). Typical algorithms that allow this are the Metropolis-Hastings MCMC algorithm (3), Hamiltonian Monte Carlo (4), or deterministic optimization methods such as Levenberg-Marquardt (5). These methods allow us to build an approximation of the posterior probability distribution over all the model parameters, from which most-likely values, confidence intervals and covariance can be extracted.

An illustration of a general Bayesian inference network is shown in **Error! Reference source not found.**. The quantities in curly braces denote sets of input parameters which are defined by their prior distributions. There are some physics model parameters and some parameters needed for the diagnostic models (e.g. calibration, alignment, etc.). To maintain generality the physics and diagnostic models are separated, but this is not necessary. Also, to maintain generality the diagnostic processing is also included in the network, as there may be a desire to estimate parameters that deal with how the data itself is processed. The diagnostic models and processed data are compared in the likelihood block. Iterating over this loop many times using a specified algorithm will result in an estimation of the posterior (on the right).

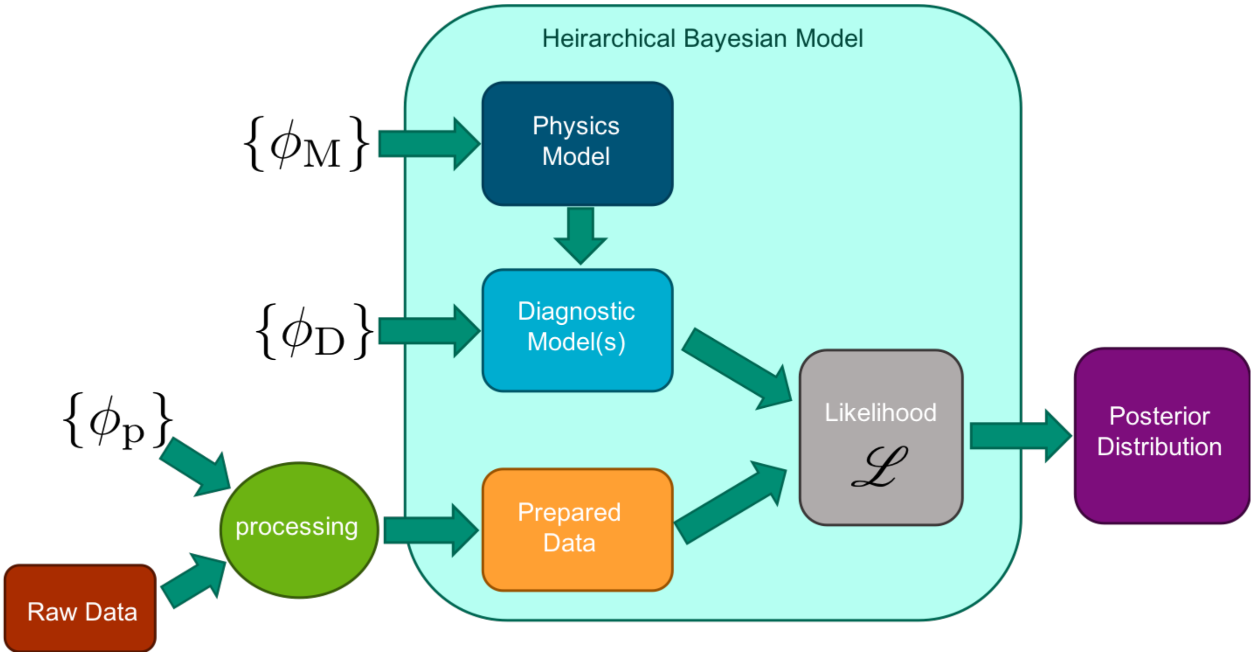

Figure 1: Illustration of a general Bayesian inference network.

One can also use this methodology to choose between competing models. This is a two step process, first the parameter estimation is performed as described above for each of the competing models. Then the evidence term is evaluated for each of these cases, which amounts to a weighting function on the goodness of fit for each of the models that accounts for the possibility of differences in complexity between them. This methodology provides a rigorous metric by which we can determine what model is most consistent with the available data. Going beyond this statement is possible as well. We can extend this formalism to ask the question: what new experiment or diagnostic should be fielded to distinguish between competing models with some confidence? Answering this requires that we have a way to quantify the information gain by adding new data relative to some baseline. This can readily be accomplished by applying metrics from information theory to the posterior obtained for each of the hypothetical cases (6) (7). Exploiting this capability can allow us to intelligently choose where to put our resources (e.g. invest in new diagnostic capabilities or lobby for experiments in a specific regime).

Unfortunately, despite the advantages and wide adoption across the physical sciences, these methods have been slow to catch on in ICF and HEDP research. This is primarily due to the computational expense of determining the posterior probability distribution as well as the conceptual difficulty of formulating one's problem in a way that is amenable to Bayesian analysis. Despite the difficulty, several independent groups across the NNSA complex have started to incorporate these methods into their work. Because of the success of these methods in other fields, the authors believe that it is imperative that we adopt them in our work. In order to speed adoption by others and solve some of the challenges we are currently facing we propose that new tools should be developed that are purpose built for our needs. Below we summarize the presentations and discussions from the workshop, followed by a recommendation of the path forward.

2. SUMMARY OF THE WORKSHOP

A workshop was held at SNL on Nov. 6th, 2019 in conjunction with the hotspot analysis workshop co-hosted by Dr. Alex Zylstra and Dr. Patrick Knapp. There were participants present from SNL, LANL, LLNL, and the LLE, with a significant number attending remotely. The purpose of this workshop was to introduce parties at each of the NNSA sites to the work being done in applying Bayesian methods to ICF and HED experiments and theory. Subsequently the discussion was based around what tools we use today, what deficiencies we see, and how we can collaborate in a way that maximizes efficiency and minimizes duplication of effort across the complex.

The morning started with talks from each of the participating sites to introduce the current scope of work at each site and frame the discussion of the path forward. Subsequently, there was a long discussion period allotted where a diverse array of topics was discussed concerning the obstacles to implementation and proliferation of these techniques. As a group, there was a clear desire to minimize duplication of effort and maximize dissemination of tools being developed. In any of these projects, a significant amount of the effort is expended in just acquiring the necessary data and putting it in a useful format. Tackling this problem in advance will speed progress for everyone. Below, we briefly summarize each of the presentations.

2.1. SNL Overview

The workshop began with a talk from Michael Glinsky articulating the path being pursued by the Bayesian methods group at SNL. This involves a paradigm where a robust Bayesian data assimilation engine sits at the heart of a loop that enables integration of theory and modeling with data to make predictions with confidence intervals and enable high quality decisions to be made. This formalism would also enable the assessment of deficiencies in models and unknown unknowns through the application of the causal statistics formalism (8) (9). Current efforts surround the construction of the Bayesian piece of this, allowing high quality inferences to be made from experimental data (10). A strong focus is on the integration of heterogeneous data to provide a view on the experiment that is consistent with all observables. SNL is also embarking on an effort to speed up the physics models, making them practical to run in these large inference loops, by creating surrogate models to approximate strategic pieces of physics and diagnostic forward models.

2.2. LLNL Overview

Jim Gaffney spoke about his work aimed at improving the predictability of ICF models by incorporating data (11). This work uses large ensembles of simulations to create a surrogate using a DNN which can be efficiently searched to match simulation to data. This step uses coarse degradation mechanisms (e.g. unspecified preheat, mix etc.) to hone in on which general mechanism is most consistent with the observables. Then a second ensemble is performed that ascribes different physical origins to the most likely degradation mechanism. For example, if preheat is found to be most consistent with the data, we want to know the origins of the anomalous preheat (e.g. M-band, hot electrons, shock mistiming, fuel ablator mix, etc.). With the uncertainties found in these inferences, trends can be examined relating degradation mechanisms to experimental input parameters, which can then be extrapolated. Uncertainties are a premium at every step along the way. Uncertainties in the DNN surrogates are modeled using the drop out method. Ongoing work will include additional diagnostics in the analysis (e.g. neutron images) as well as other HED experiments that can help constrain degradation mechanisms and physical models.

2.3. LLE Overview

Varchas Gopalaswamy spoke on two parallel efforts at the LLE to 1) use calibrated models to inform experimental design and scaling and 2) analyze experimental data. The main focus of item 1 is to use the 1D code LILAC (12) and experimental data to map 1D predictions to experimental outcomes (13). Log-linear regression is what has been primarily used to construct the fast model with more recent work using Relu back-propagation neural networks for the regression. This fast model can then be used to propose new promising experimental designs, which can be checked via LILAC simulations and eventually experiments if they bear out. There is also an effort to use Bayesian methods to analyze experimental data. A complex method is being considered to model the experiment in LILAC up to peak velocity and then apply perturbations and evolve the hotspot in 3D, comparing the synthetic observables to experimental data. A simpler method is being employed to constrain the 3D hotspot and shell shape at bang time using multiple lines of sight. Finally, there is also an effort to use multiple complimentary shots (layered cryo shots plus scaled CD shots) to constrain the hot electron deposition in the cold fuel layer in the cryo shots. This work is ongoing, but shows promise in relating the predicted and measured hard x-ray signals.

2.4. LANL Overview

LANL is approaching the application of these methods from the measurement science perspective. LANL is using the BIE to infer parameters from radiography experiments on OMEGA and the NIF (14). This method uses a ray tracing algorithm to remove bias from the experiments by inferring the best alignment, magnification, and other experimental setup parameters that fit the observations best. Once these experimental parameters are accounted for, the algorithm can be used to infer quantities of interest (e.g. shock position, tracer layer width, etc.) with very high confidence. A feature of the BIE that is particularly useful is that each module must have a gradient/adjoint specified, allowing acceleration of optimization and facilitating sensitivity studies. The driving motivations for employing these methods is to make more of the data collected and make the most of the allocated experimental time. This is done by improving the understanding of the diagnostics and experimental configuration, performing sensitivity studies to maximize the utility of the diagnostics and experiments, and by emphasizing statistical significance of the comparison between models and data. LANL is also pursuing the use of machine learning to process data (e.g. denoising, feature detection). There was an emphasis that synthetic diagnostics play a central role in statistical inference, data processing, and experimental design.

2.5. Comments

As seen in each of the presentations, the four sites represented at the workshop are tackling different problems and each are using different tools in order to accomplish this. LLNL is heavily leveraging machine learning and their infrastructure that has been developed to run large ensembles of rad-hydro simulations. The other sites tend towards an approach using simpler models of the experiments. All are interested in ways to combine diagnostics and improve confidence in our inferences.

3. PATH FORWARD

As evidenced by the presentations and discussion there is strong interest in implementing Data Science methods to improve the use of our data and facilitate the integration of theory and experiment. However, exactly how to do this is somewhat challenging. Below we outline the challenges that currently exist and how we might overcome them.

3.1. Data Standardization and Management

A significant amount of effort is expended every time a new project is undertaken simply acquiring and organizing all of the data needed to perform the task. Because high quality diagnostic models are needed to perform Bayesian inference, information about the instruments (e.g. calibration, characterization, geometry) must be accessible. Additionally, the data itself must be accessible as well as the necessary experimental setup information. Often, these data are stored in different places, with different formats, and there may be no way to connect certain pieces of information with specific experiments. It is quite common for the same type of data to be kept in excel spreadsheets of various structures with the meta information contained in the directory structure and file name. For these cases, it is necessary to manually mine the different systems which often requires significant time and effort, and it is inherently error prone.

The NIF shot archive is a good example of how this can be done better. An ideal case would connect experiment ID to data, diagnostic configurations, calibrations and up to date diagnostic models, target data, experimental configuration, and relevant simulations that were performed in support of the experiment. Seamless integration of each of these archives would facilitate post-shot simulations with accurate synthetic diagnostics as well as analysis of the data acquired on the experiment with high fidelity forward diagnostic models.

Frameworks exist that allow for the storing and access of unstructured, heterogeneous data. Choosing a data standard that multiple sites can adopt so that tools that utilize the data can be shared and developed across the sites would speed adoption. In order for this to be effective there must be agreed upon standards for data management, inclusion of metadata, version control, and the API that facilitates access. We should examine the success and shortcomings of the NIF shot archive and develop a tool that can be robust, flexible, and widely adopted.

3.2. Common Tools for Model Building and Inference

There are a number of tools in existence that facilitate Bayesian inference, many of which are python based making them widely used. However, each of the general purpose libraries have significant shortcomings. LANL uses a tool called the Bayesian Inference Engine (BIE) that uses a LabView-like building block structure to construct the inference network. The user can create custom blocks to perform their specific task. One of the aspects that is highly desirable is that each of the blocks has a derivative/adjoint defined which can significantly improve the efficiency of estimating the posterior. However, the BIE is written in SmallTalk which limits the user base and does not natively compute confidence intervals. SNL and LLNL use an open source python library called pyMC (15). This package has a number of nice features. By programming one's model using the building blocks of the pyMC package, the software creates a hierarchical graph to define the inference network. However, it quickly becomes very complex when dealing with realistic physics and diagnostic models that are not built using the native tools of pyMC. Sandia implemented several extensions to the pyMC2 package to overcome these issues and include additional optimization algorithms. Additionally, the support for the use of gradients is not as deeply engrained in the software making certain algorithms difficult or impossible to use. Finally, most of these packages have poor support for optimization methods that are robust to nonconvex surfaces or are parallelizable.

When faced with this situation, it is common to settle for a tool that does not have all of the necessary features and either work within its limitations or expend significant effort to extend its capabilities. However, given the required level of effort to get these tools to a useful state it is perhaps a better use of time to develop one that already has the necessary components. A new tool should satisfy the following criteria:

- Interface easily with the data storage architecture
- Use python or other widely accessible high level language as the primary user interface
- Facilitate the construction of complex inference networks with minimal development overhead
- Facilitate the use of gradients in all aspects and support a wide variety of global and local optimization techniques
- Support the use of modern machine learning tools (e.g. TensorFlow, pyTorch, etc.)
- Integrate natively with the labs' HPC resources
- Provide a framework for creating physics and diagnostic models that facilitates rapid development and integration

The final point is an important one. With existing packages, a significant amount of time is spent figuring out how to retrofit existing models or write new ones in such a way that is compatible with the API for the Bayesian machinery. We believe the package should have an API expressly for this purpose and should also contain a library of primitives that are already built to interface directly with the machinery and take advantage of its features.

3.3. Towards Analysis Standards for ICF and HED Research

An overarching goal of this discussion is to enable researchers in ICF and HED to do better analysis more efficiently. This will, in turn, enable more confident integration of theory and experiment, which drives progress in our knowledge. We often find ourselves in the position where data disagree with simulations and it is not obvious which source of information is more reliable. We know that our models are incomplete, but we also know that our understanding of our experimental configurations and diagnostics is incomplete as well. This situation can cause progress to slow, or even stop, as we are unable to use data to change our models. By implementing a fully Bayesian approach to data analysis that accounts for these uncertainties we will then be able to determine exactly how confident we are in our data. If these approaches are adopted broadly across the complex, we then have a common language and methodology with which to make these assessments.

We suggest that, in addition to implementing standards for data management and developing a tool for Bayesian inference that can be used broadly, we should also be developing standards for data analysis that address these concerns. We should be establishing methodologies for modeling instruments and experiments that maximize our confidence in measurements, thereby maximizing our ability to use measurement to advance theory and modeling.

4. CONCLUSIONS

This White Paper is a call to action for the funding and development of a Data Science infrastructure for ICF and HEDP. This must be developed with a full use case (that is real HEDP and ICF experiments) as the test problems for the development. Furthermore, both designers and experimentalists should be part of the development team, along with data scientists and computer scientists.

This infrastructure will have two major components. The first is data stewardship. An infrastructure must be developed for both experimental and simulation data that is: 1) consistent and standardized, 2) self-descriptive, 3) efficient and parallelized, 4) based on open formats, and 5) incorporates uncertainty. The second is an assimilation engine. This engine has four major parts: 1) a Bayesian inference network, that is a graph where the nodes are models and the edges are physical quantities, 2) a sampling engine of the distribution generated from the Bayesian inference network, 3) fast surrogates of the physics, both rad-MHD simulations and synthetic experimental diagnostics, probably involving physics constrained MLDL, and 4) model selection & proposal, based on causal statistics.

The sampling engine must be complete. There are 12 key ingredients that it must have. They are: 1) multi-model mixture models (e.g., genetic algorithms, multi-start optimization, and EM algorithm), 2) sampling efficiency (e.g., importance sampling and LM optimization), 3) nonlinear distributions (e.g. Metropolis algorithm), 4) boundaries (e.g., projected Newton methods), 5) stiff objective sampling (e.g., Bayesian bootstrap), 6) prior bias correction (e.g., Bayesian whisper iteration), 7) visualization, 8) value of information including sensitivity output and visualization, 9) parallelization of computation, 10) analytic adjoints of forward models, 11) multi-fidelity models (i.e., how to combine and determine computation budget for each fidelity), and 12) software architecture standardization (i.e., physics model, diagnostics, etc.).

This is a significant effort, whose success is dependent on a multi-disciplinary team that understands data science, computer science, the physics-based models, and the experiment. The payoff will be in better decisions, informed by the data science. These decisions range from the strategic, things like infrastructure investment and diagnostic development, to the tactical, things like experimental design and scientific hypothesis confirmation.

REFERENCES

1. **Sivia, D.** *Data Analysis: A Bayesian Tutorial.* second. s.l. : Oxford Science Publications, 2006.
2. *Bayesian inference in physics.* **Toussaint, Udo Von.** July-September 2011, Reviews of Modern Physics, Vol. 83.
3. *Equation of State by fast coputing machines.* **Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N. Teller, A.H., Teller, E.** 1953, Journal of Chemical Physics, Vol. 21, pp. 1087-92.
4. **Michael Betancourt, Mark Girolami.** Hamiltonian Monte Carlo Methods for Hierarchical Models. [book auth.] Umesh Singh, Dipak K. Dey, Appaia Loganathan Satyanshu K. Upadhyay. *Current Trends in Bayesian Methodology with Applications.* s.l. : CRC Press, 2015.
5. **More, J.J.** The Levenberg-Marquardt algorithm: Implementation and Theory. [book auth.] Watson G.A. *Numerical Analysis.* s.l. : Springer, 1978.
6. **Shannon, C. E.** *A Mathematical Tehory of Communication.* Bell Labs. s.l. : Bell System Technical Journal, 1948. pp. 379-423, 623-656.
7. *On Information and Sufficiency.* **Leibler, Solomon Kullback and Richard.** 79-86, 1951, The Annals of Mathematicalk Statistics, Vol. 22.
8. **Pearl, Judea.** *Causality: Models, Reasoning and Inferen e.* s.l. : Cambridge University Press, 2009.
9. *Causal Discovery and Inference: Concepts and Recent Methodological Advances.* **Zhang, Peter Spirtes and Kun.** 3, s.l. : Springer, 2016, Applied Informatics, Vol. 3.
10. *Bayesian Parameter Estimation for Inertial Confinement Fusion Experiments Using Multi-Objective Optimization.* **P.F. Knapp, et al.** In preparation.
11. *Making Inertial Confinement Fusion Models More Predictive.* **Gaffney, J. A. et al.** s.l. : AIP, 2019, Phyiscs of Plasmas, Vol. 26.
12. *Effect of laser illumination nonuniformity on the analysis of time-resolved X-ray measurements in UV spherical transport experiments.* **Delettrez, J., Epstein, R., Richardson, M. C., Jaanimagi, P. A. & Henke, B. L.** 1987, Physical Review A, Vol. 36, pp. 3926-3934.
13. *Tripled Yield in direct-drive laser fusion through statistial modeling.* **V. Gopalaswamy, et al.** 565, 2019, Nature, pp. 581-586.
14. *Demonstration of Scale-Invariant Rayleight-Taylor Instability Growth in Laser-Drive Cylindrical Implosion Experiments.* **J.P. Sauppe, et al.** 124, 2020, Physical Review Letters, p. 185003.
15. *Probabilistic programming in Python using PyMC3.* **Salvatier J., Wiecki T.V., Fonnesbeck C.** 2016, PeerJ Computer Science .
16. *Predicting the output from a complex computer code when fast approximations are available.* **O'Hagan, M.C. Kennedy and A.** s.l. : Biometrika, 2000, Vol. 87.